Bayesian UQ for subsurface inversion using multiscale hierarchical model

Bani Mallick, Yalchin Efendiev, Akhil Datta-Gupta, Anirban Mondal

bmallick@stat.tamu.edu

Department of Statistics, Texas A&M University, College Station

Bayesian analysis and UQ

Goal is to estimate:

- Model parameters and their **uncertainties.**
- Predictive uncertainty distribution for future responses.

Bayesian approach to analysis:

- Focus is on uncertainties in parameters, as much as on their best (estimated) value.
- Permits use of prior knowledge, e.g., previous experiments, modeling expertise, physics constraints.
- Model-based.
- Can add data sequentially

Data Integration

- Bayesian hierarchical models are natural tools for combining information from diverse sources
- Data at different scales and spatially correlated: reservoir data
- Data from different studies and borrow information across both subjects and studies
- Data could be from cross-platform
- Gene expression data: Depending on the technology the expression data it can be continuous (microarray) or discrete (SAGE, MPSS)

Forward Model and Inverse problem

$$Z = F(\tau) + \epsilon$$

where

- F is the forward model, simulator, computer code which is non-linear and expensive to run.
- τ input parameter: could be of very high dimension.
- Z is the observed response.
- ϵ is the random error usually assumed to be Gaussian.
- Want to estimate τ with UQ.
- This is a non-linear inverse problem.

Fluid flow in porous media

- Studying flow of liquids (Ground water, oil) in aquifer (reservoir).
- Applications: Contaminant cleanup, Oil production.
- Forward Model:Flow of liquid (or production data, output) when the physical characteristics (permeability, porosity) are known.
- Inverse problem: Inferring the permeability (porosity) from flow data.

Permeability

- Primary parameter of interest is the permeability field.
- Permeability is a measure of how easily liquid flows through the aquifer at that point.
- This permeability values vary over space.

Forward Model

Darcy's law:

$$v_j = -\frac{k_{rj}(S)}{\mu_j} k_f \nabla p, \tag{1}$$

- v_j is the phase velocity
- k_f is the fine-scale permeability field
- k_{rj} is the relative permeability to phase j (j=oil or water)
- S is the water saturation (volume fraction)
- p is the pressure.

Forward Model

Combining Darcy's law with a statement of conservation of mass allows us to express the governing equations in terms of pressure and saturation equations:

$$\nabla \cdot (\lambda(S) k_f \nabla p) = Q_s, \tag{2}$$

$$\frac{\partial S}{\partial t} + v \cdot \nabla f(S) = 0, \tag{3}$$

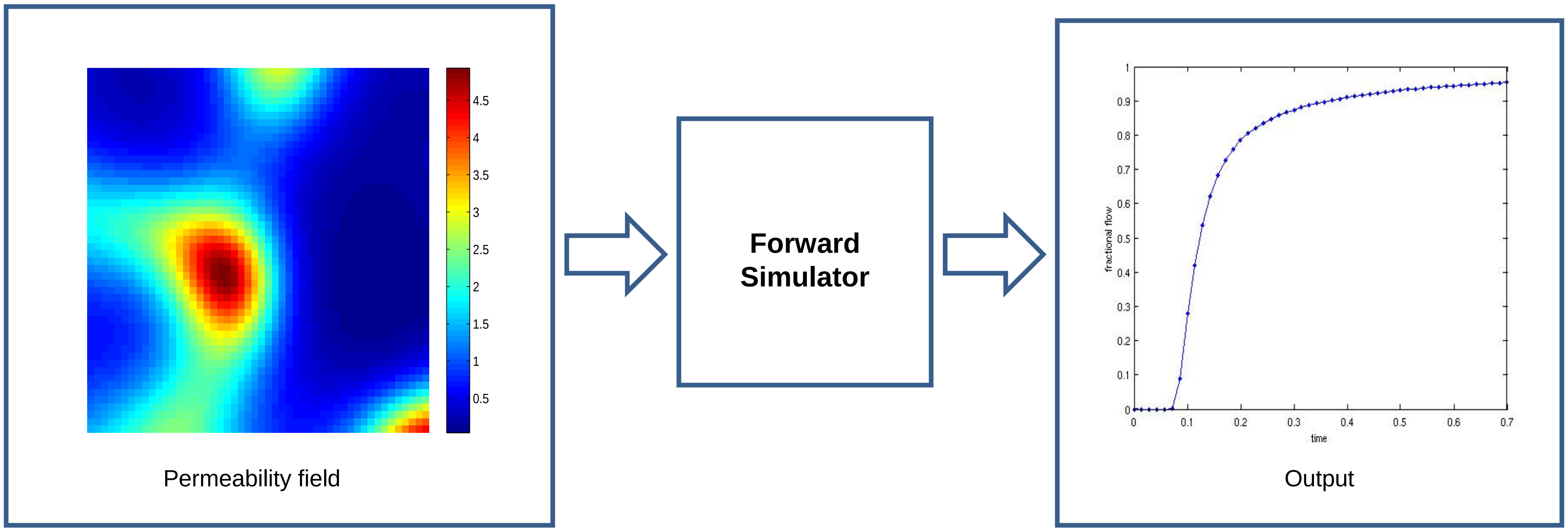
- λ is the total mobility
- Q_s is a source term
- f is the fractional flux of water
- v is the total velocity

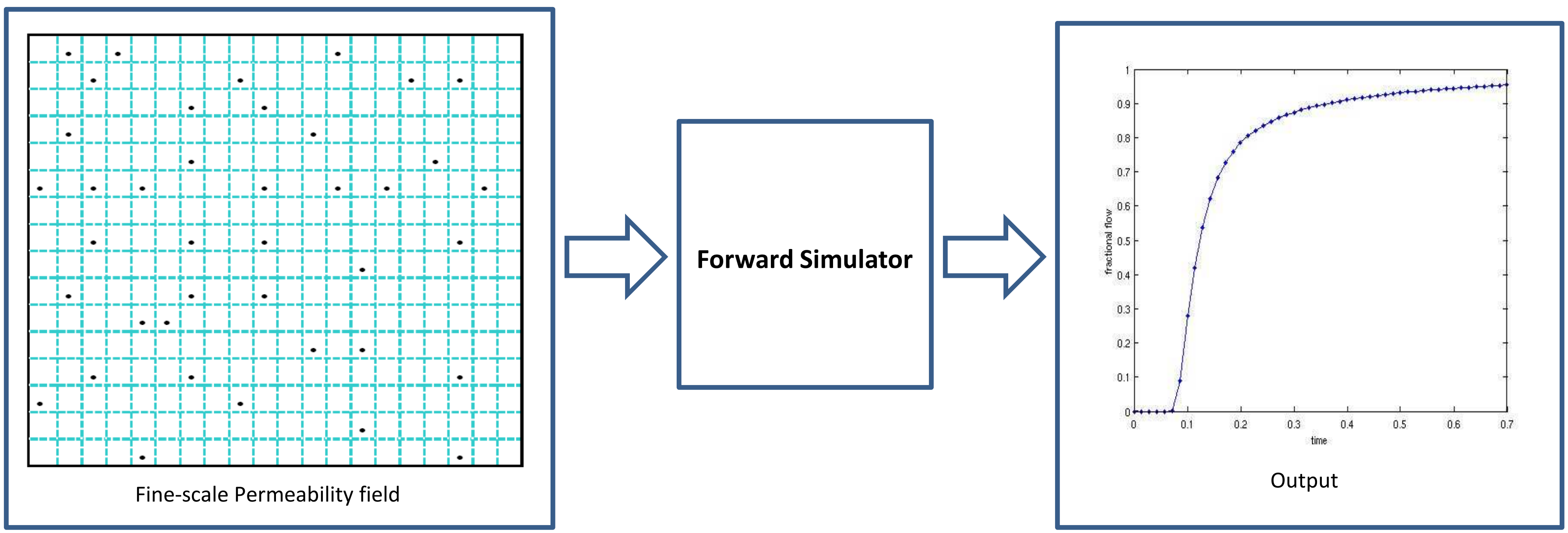
Forward Model

Production (amount of oil in the produced fluid, fractional Flow or water-cut) $F(k_f)$ is given by

$$F(k_f) = \int_{\partial\Omega^{out}} v_n f(S) dl$$

where $\partial\Omega^{out}$ is outflow boundaries and v_n is normal velocity field.


4.5
4
3.5
3
2.5
2
1.5
1
0.5
Permeability field
Forward Simulator
fractional flow
time
Output



Fine-scale Permeability field
Forward Simulator
fractional flow
time
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Output

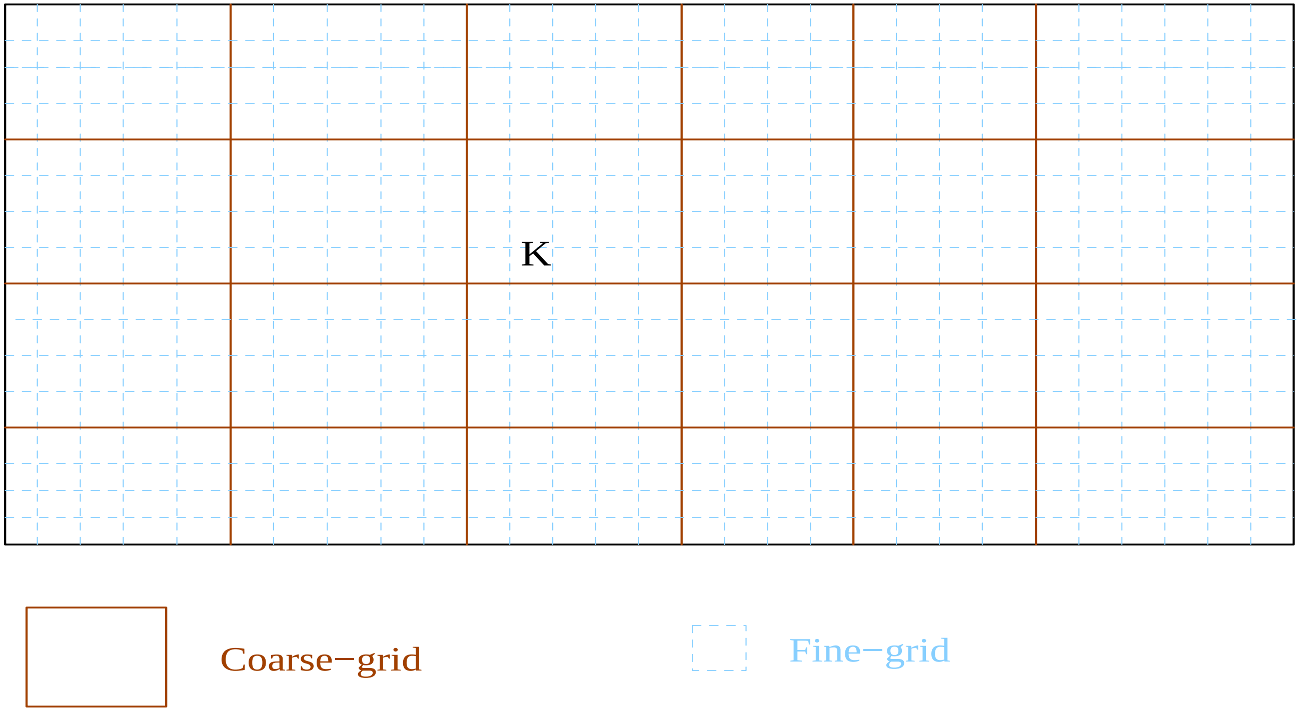
Inverse Problem

$$Z = F(k_f) + \epsilon$$

- Z is the observed production data.
- F is the forward simulator which is the solution of a coupled nonlinear pde's.
- k_f is the fine-scale permeability field of high dimension.
- ϵ is the random error.

Forward Model

- We want to infer k_f conditioned on Z (history matching)
- Some observed fine-scale permeability values (k_f^0) are available but expensive (well logs, cores)
- Additional data: coarse-scale permeability data (k_c) from seismic traces
- We want to model the fine scale permeability field condition on the observe fine scale data, coarse scale data and the production data.



No flow

$\phi = 1$ $$div(k_f(x)\Delta\phi) = 0$$ $\phi = 0$

No flow

$$(k_c(x)e_j, e_l) = \frac{1}{|K|}\int_K (k_f(x)\Delta\phi_j(x), e_l)dx$$

Bayesian Framework

- Number of parameters in the permeability field is large relative to the number of available data points.
- Dimension reduction:Replacing k_f by a finite set of parameters τ.
- Building enough structures through models and priors.
- Need to link data at different scales.
- Bayesian hierarchical models have the ability to do all these things simultaneously.

Bayesian Framework

Bayesian model: Joint probability specification on Data: Z and unknown τ through $P(Z, \tau)$.
The convenient way to express it:

$$P(Z, \tau) = P(Z|\tau)P(\tau)$$

- $P(\tau)$: Prior density of τ.Ex: Non-informative prior, Priors based on physical principles (positivity, invariance arguments), Priors based on previous experiments, Prior from expert opinions.
- $P(Z|\tau)$: Likelihood function: Conditional density of $Z|\tau$: Gaussian model is popular one. Heavy tailed distributions to accommodate outliers. In our model distribution of ϵ controls it.

Likelihood calculations

$$Z = F(\tau) + \epsilon$$

For Gaussian model the likelihood will be

$$P(Z|\tau) = \frac{1}{\sqrt{2\pi}\sigma_1} Exp(\frac{-[Z - F(\tau)]^2}{2\sigma_1^2})$$

where σ_1^2 is the variance of ϵ.

Likelihood Calculations

- It is like a black-box likelihood which we can't write analytically, although we do have a code F that will compute it.
- We need to run F to compute the likelihood which is expensive.
- Hence, no hope of having any conjugacy in the model, other than for the error variance in the likelihood.
- Need to be somewhat intelligent about the update steps during MCMC so that do not spend too much time computing likelihoods for poor candidates.

Posterior Density

Posterior density of τ: $P(\tau|Z)$ [Uncertainty of τ after observing the data Z]

$$P(\tau|Z) = \frac{P(\tau)P(Z|\tau)}{P(Z)}.$$

- Posterior Density provides the uncertainty distribution of the unknown parameters.
- Provides complete quantitative description of uncertainties.

Prediction and MCMC

$$P(Z_{new}|Z_{obs}) = \int_{\tau} P(Z_{new}|Z_{obs},\tau)P(\tau|Z_{obs})d\tau$$

- For complex, nonlinear models, posterior will be not in explicit form.
- Simulate samples of the parameters from the posterior distribution rather than explicit solution.
- These samples will be utilized to construct the posterior uncertainty distribution of the parameters.
- High dimensional parameter space, hence we use Markov chain Monte Carlo method (MCMC).
- These samples can be used to perform Monte Carlo integration to obtain the predictive distribution.

Procedures

- Reduce the dimension of the permeability field.
- Use the reduced dimension parameters as input parameters τ.
- Use MCMC to draw samples from $P(\tau|Z)$.
- Avoid repeated calculations of the expensive likelihood using two stage MCMC.

Dimension reduction

- We need to reduce the dimension of the permeability field K_f
- This is a spatial field denoted by $K_f(\boldsymbol{x}, \omega)$ where $\boldsymbol{x}$ is for the spatial locations and ω denotes the randomness in the process
- Assuming K_f to be a real-valued random field with finite second moments we can represent it by Kauren-Loeve (K-L) expansion

K-L expansion

$$K_f(\boldsymbol{x}, \omega) = \theta_0 + \sum_{l=1}^{\infty} \sqrt{\lambda_l}\theta_l(\omega)\phi_l(\boldsymbol{x})$$

where

- λ: eigen values
- $\phi(\boldsymbol{x})$ eigen functions
- θ: uncorrelated with zero mean and unit variance
- If K_f is Gaussian process then θ will be Gaussian

K-L expansion

If the covariance kernel is C then we obtain them by solving

$$\int C(\boldsymbol{x}_1, \boldsymbol{x}_2)\phi_l(\boldsymbol{x}_2)d\boldsymbol{x}_2 = \lambda_l \phi_l(\boldsymbol{x}_1)$$

and can express C as

$$C(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{l=1}^{\infty} \lambda_l \phi_l(\boldsymbol{x}_1)\phi_l(\boldsymbol{x}_2)$$

Spatial covariance

We assume the correlation structure

$$C(\mathbf{x},\mathbf{y}) = \sigma^2 \exp\left(-\frac{|x_1-y_1|^2}{2l_1^2} - \frac{|x_2-y_2|^2}{2l_2^2}\right).$$

where, l_1 and l_2 are correlation lengths.
For an m-term KLE approximation

$$\begin{aligned} K_f^m &= \theta_0 + \sum_{i=1}^{m} \sqrt{\lambda_i}\theta_i\Phi_i, \\ &= B(l_1, l_2, \sigma^2)\theta, \text{(say)} \end{aligned} \quad (1)$$

Existing methods

- The energy ratio of the approximation is given by

$$e(m) := \frac{E\|k_f^m\|^2}{E\|k_f\|^2} = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^\infty \lambda_i}.$$

- Assume correlation length l_1, l_2 and σ^2 are known.
- We treat all of them as model parameters, hence $\tau = (\theta, \sigma^2, l_1, l_2, m)$.

Inverse Problem

- We want to infer k_f conditioned on Z.
- Additional data: coarse-scale permeability field k_c.
- Some of the observed fine-scale permeability values k_f^o, at the well locations.

Hierarchical Bayes' model

$$P(\theta, l_1, l_2, \sigma^2 | Z, k_c, k_f^o) \propto P(z|\theta, l_1, l_2, \sigma^2) P(k_c|\theta, l_1, l_2, \sigma^2)$$
$$P(k_f^o|\theta, l_1, l_2, \sigma^2) P(\theta) P(l_1, l_2) P(\sigma^2)$$

- $P(z|\theta, l_1, l_2, \sigma^2)$: Likelihood
- $P(k_c|\theta, l_1, l_2, \sigma^2)$: Upscale model linking fine and coarse scales
- $P(k_f^o|\theta, l_1, l_2, \sigma^2)$: Observed fine scale model
- $P(\theta)P(l_1, l_2)P(\sigma^2)$: Priors

Likelihood

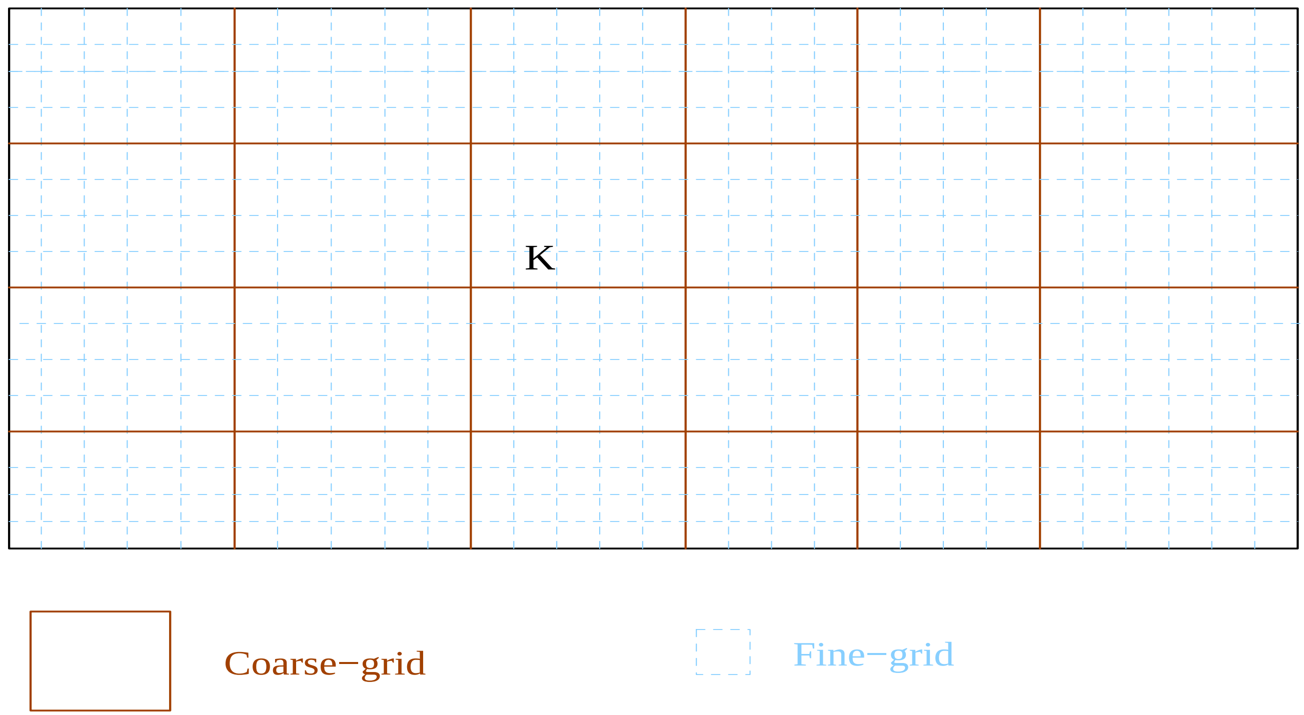
The likelihood can be written as follows:

$$
\begin{aligned}
Z &= F[B(l_1, l_2, \sigma^2)\theta] + \epsilon_f \\
&= F_1(\theta, l_1, l_2, \sigma^2) + \epsilon_f
\end{aligned}
$$

where, $\epsilon_f \sim MVN(0, \sigma_f^2 I)$.

Coarse model and upscaling

- Upscaling technique to obtain the coarse models from the fine model.
- For coarsening the fine-scale permeability field we use single-phase flow upscaling procedure for two-phase flow in heterogeneous porous media.
- The main idea of this approach is to upscale the absolute permeability field k on the coarse-grid, then solve the original system on the coarse-grid with upscaled permeability field.
- The calculation of a coarse-scale permeability is that it delivers the same average response as that of the underlying fine-scale problem locally.



No flow

$\phi = 1$ $$div(k_f(x)\Delta\phi) = 0$$ $\phi = 0$

No flow

$$(k_c(x)e_j, e_l) = \frac{1}{|K|}\int_K (k_f(x)\Delta\phi_j(x), e_l)dx$$

Upscale model

The Coarse-scale model can be written as follows.

$$\begin{aligned} k_c &= L_1(k_f) + \epsilon_c \\ &= L_1(\theta, l_1, l_2, \sigma^2) + \epsilon_c \end{aligned}$$

where, $\epsilon_c \sim MVN(0, \sigma_c^2 I)$.
i.e $k_c|\theta, l_1, l_2, \sigma^2, \sigma_c^2 \sim MVN(L_1(\theta, l_1, l_2, \sigma^2), \sigma_c^2 I)$.

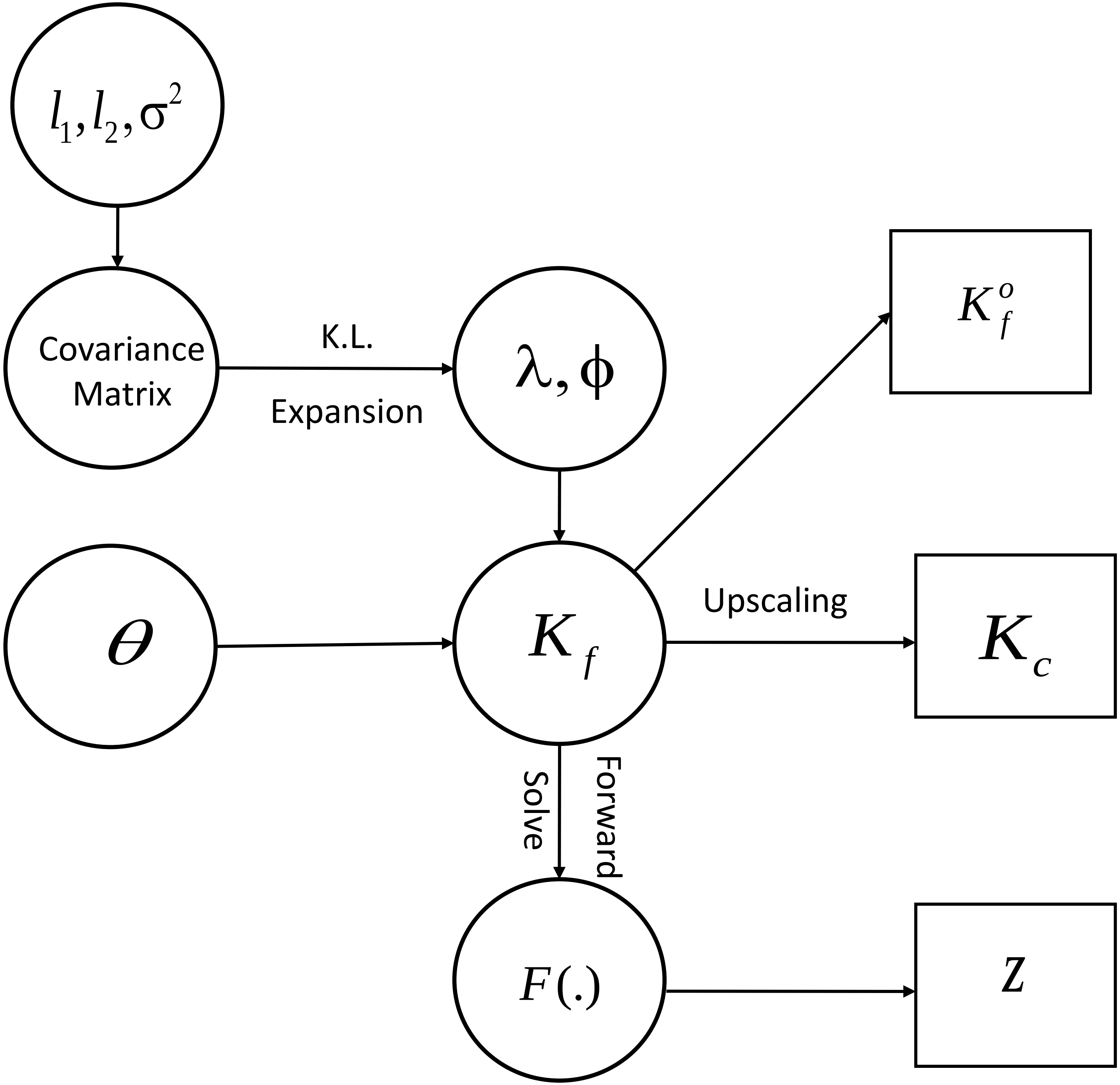
Observed fine scale model

We assume the model $k_f^o = k_p^o + \epsilon_k$

where, $\epsilon_k \sim MVN(0, \sigma_k^2)$.

k_p^o is the permeability-field obtained from K-L the expansion at the observed well locations.

So here we assume, $k_f^o | \theta, l_1, l_2, \sigma^2, \sigma_k^2 \sim MVN(k_p^o, \sigma_k^2)$,



l_1, l_2, σ^2
Covariance Matrix
K.L.
Expansion
λ, ϕ
θ
K_f
K_f^o
Upscaling
K_c
Solve
Forward
$F(.)$
Z

Inverse problem

- We can show that the posterior measure is Lipschitz continuous with respect to the data in the total variation distance
- It guaranties that this Bayesian inverse problem is well-posed
- Say, y is the total dataset, i.e, $y = \begin{pmatrix} z \\ k_c \\ k_f^0 \end{pmatrix}$
- $g(\tau, y)$ is the likelihood and $\pi_0(\tau)$ is the prior

Inverse problem

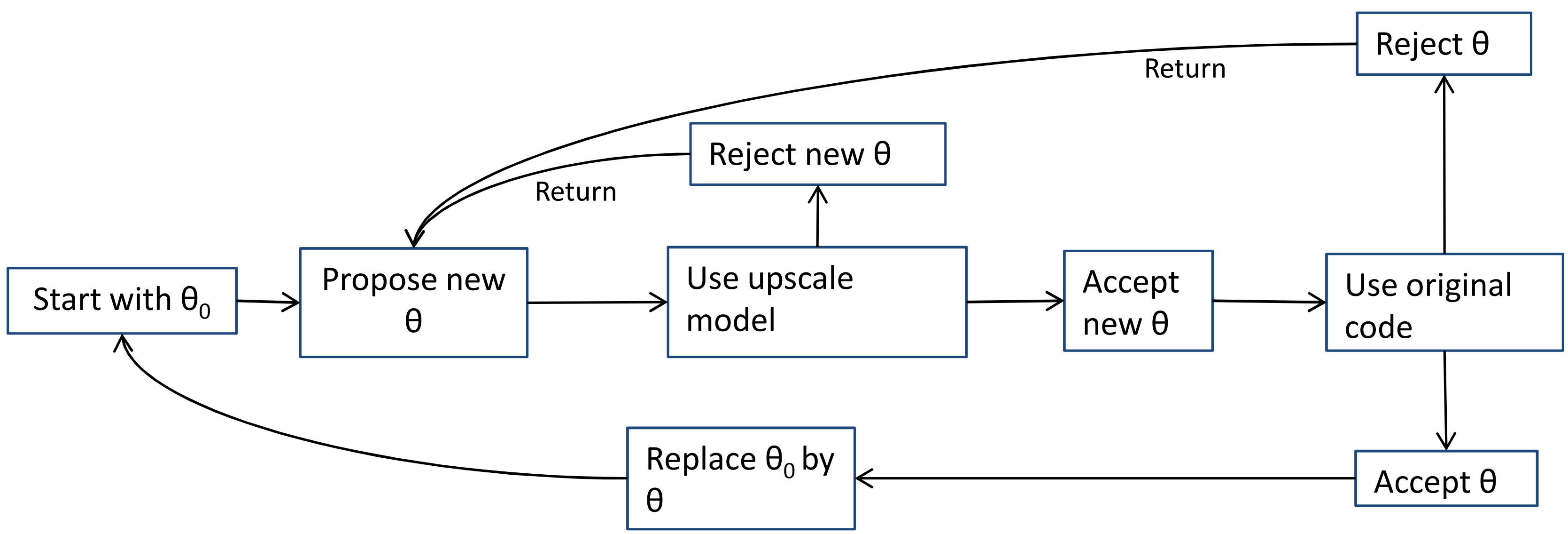
Theorem 0.1. $\forall\, r > 0,\ \exists\, C = C(r)$ *such that the posterior measures* π_1 *and* π_2 *for two different data sets* y_1 *and* y_2 *with* $max\,(\|y_1\|_{l^2}, \|y_2\|_{l^2}) \leq r$*, satisfy*

$$\|\pi_1 - \pi_2\|_{TV} \leq C\|y_1 - y_2\|_{l_2},$$

MCMC computation

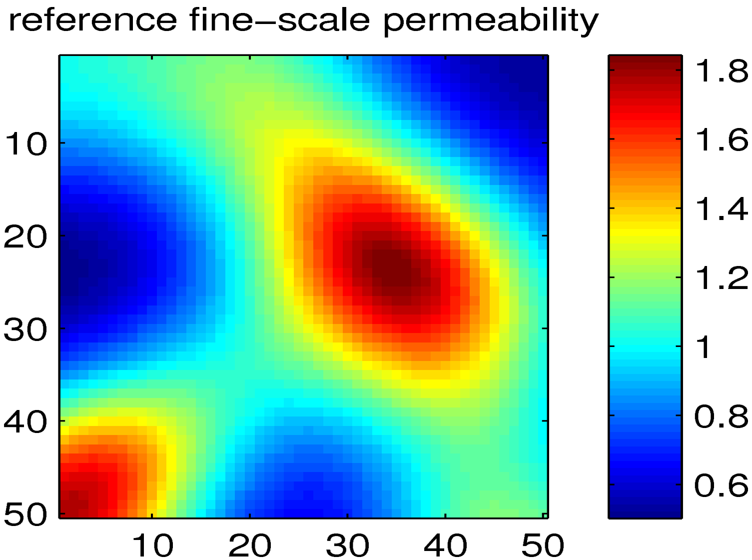
- Metropolis-Hastings (M-H) Algorithm to generate the parameters.
- Reversible jump M-H algorithm when the dimension m of the K-L expansion is treated as model unknown.
- Two step MCMC or Langevin can accelerate our computation.

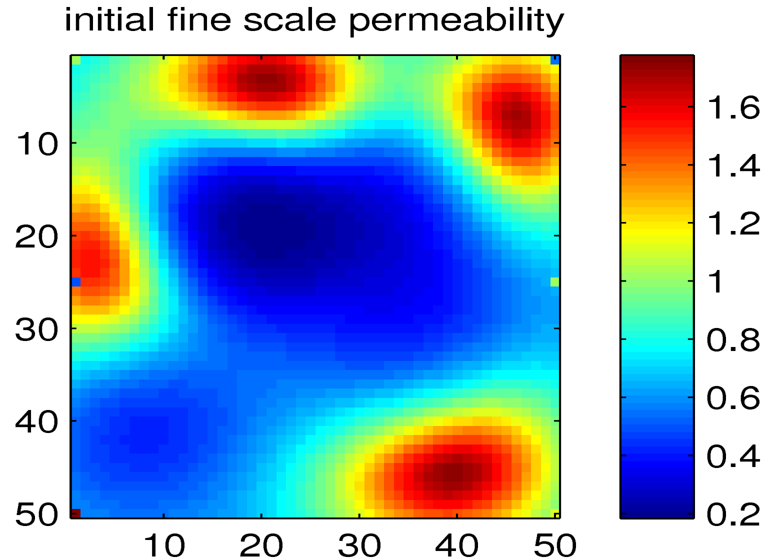
- Two stage Metropolis

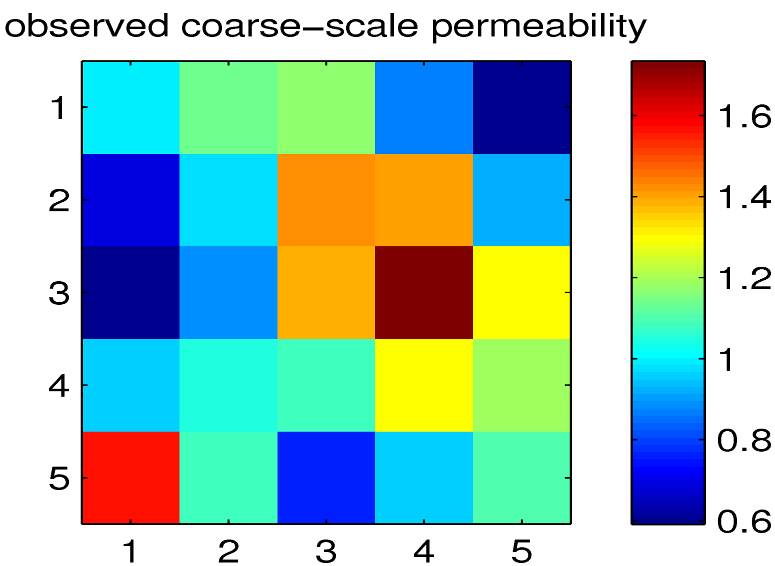


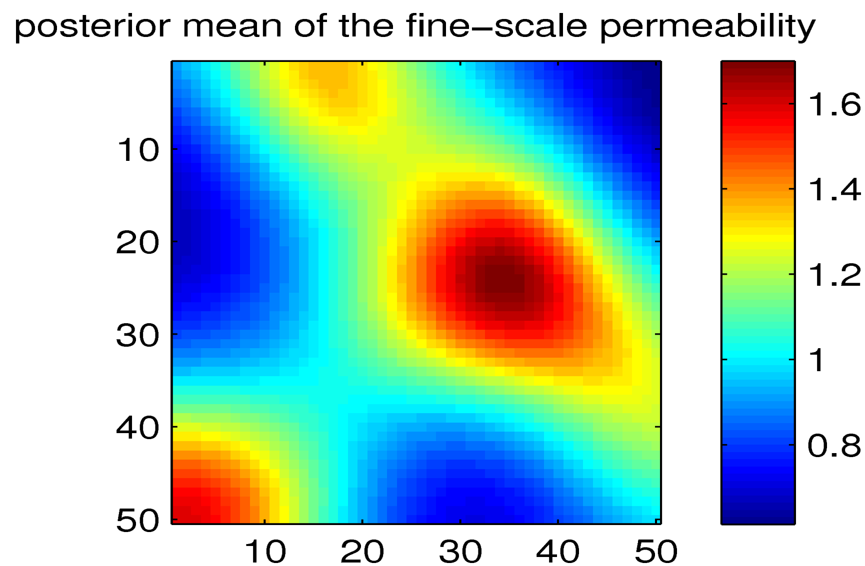
Numerical Results

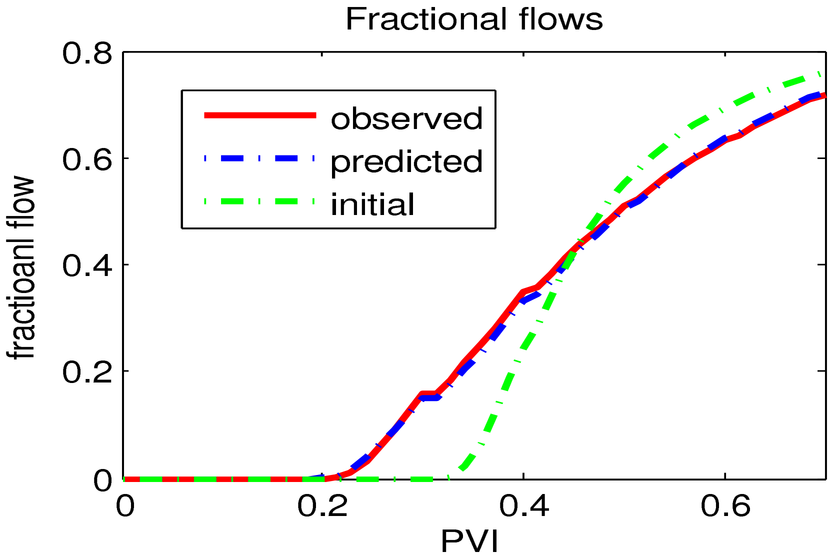
- In our first example we have considered only the isotropic case, i.e we take $l_1 = l_2 = l$, (say)
- We consider a 50×50 fine-scale permeability field on unit square.
- We generate 15 fine-scale permeability field with $l = .25$, $\sigma^2 = 1$ and the reference permeability field is taken to be the average of these 15 permeability field.
- The observed coarse-scale permeability field is calculated using the upscaling procedure in a 5×5 coarse grid.
- First, we analyzed when 10% fine-scale data are observed with the coarse scale data

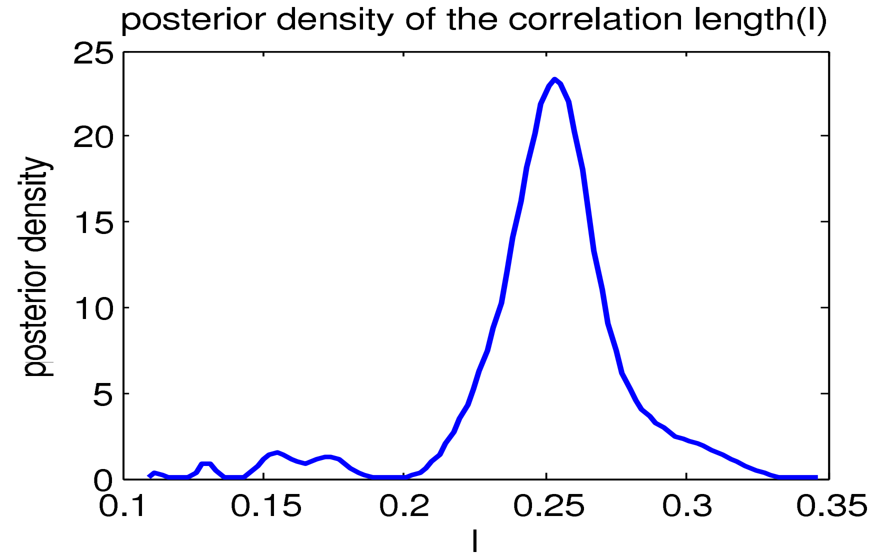

reference fine−scale permeability
10
20
30
40
50
10 20 30 40 50
1.8
1.6
1.4
1.2
1
0.8
0.6

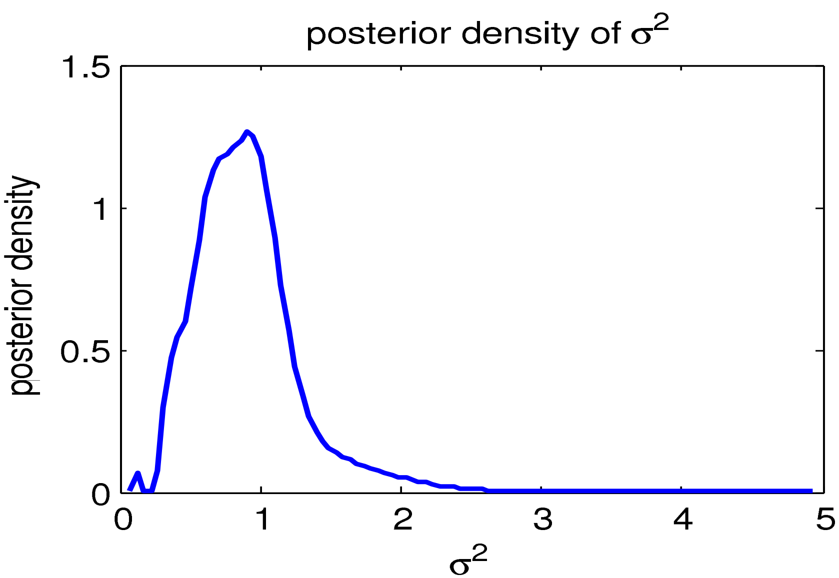

initial fine scale permeability
10
20
30
40
50
10 20 30 40 50
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2

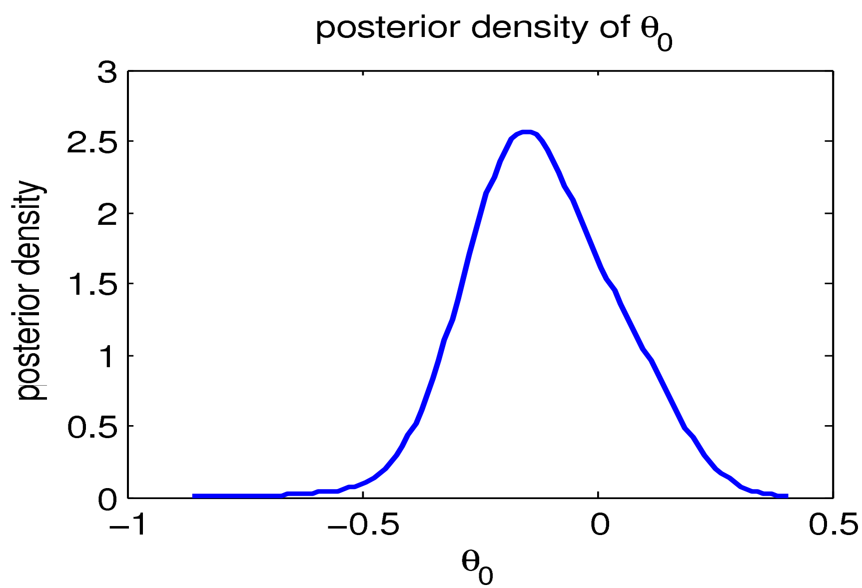

observed coarse−scale permeability
1
2
3
4
5
1 2 3 4 5
1.6
1.4
1.2
1
0.8
0.6

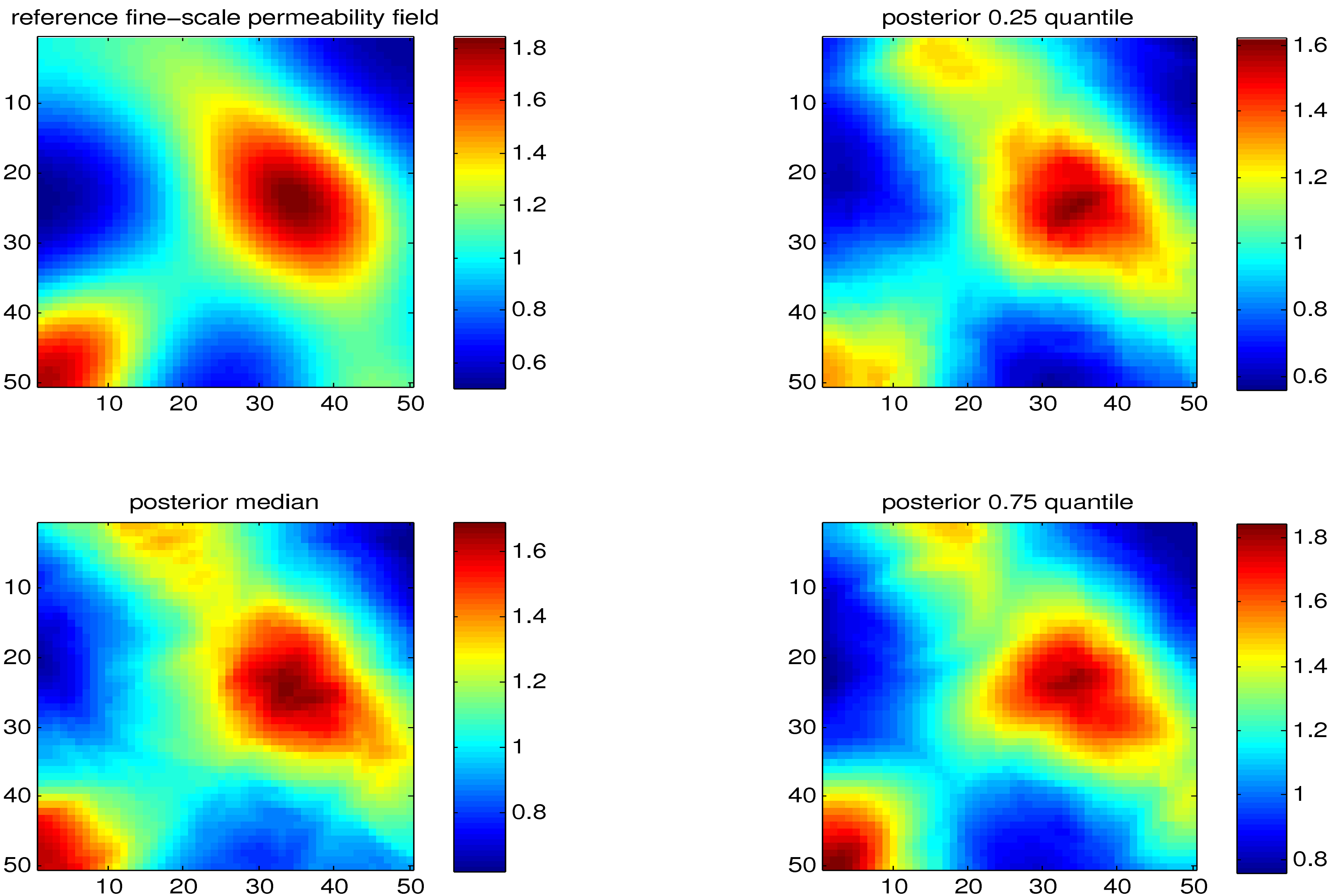

posterior mean of the fine−scale permeability
10
20
30
40
50
10 20 30 40 50
1.6
1.4
1.2
1
0.8


Fractional flows
fractioanl flow
PVI
observed
predicted
initial
0
0.2
0.4
0.6
0.8

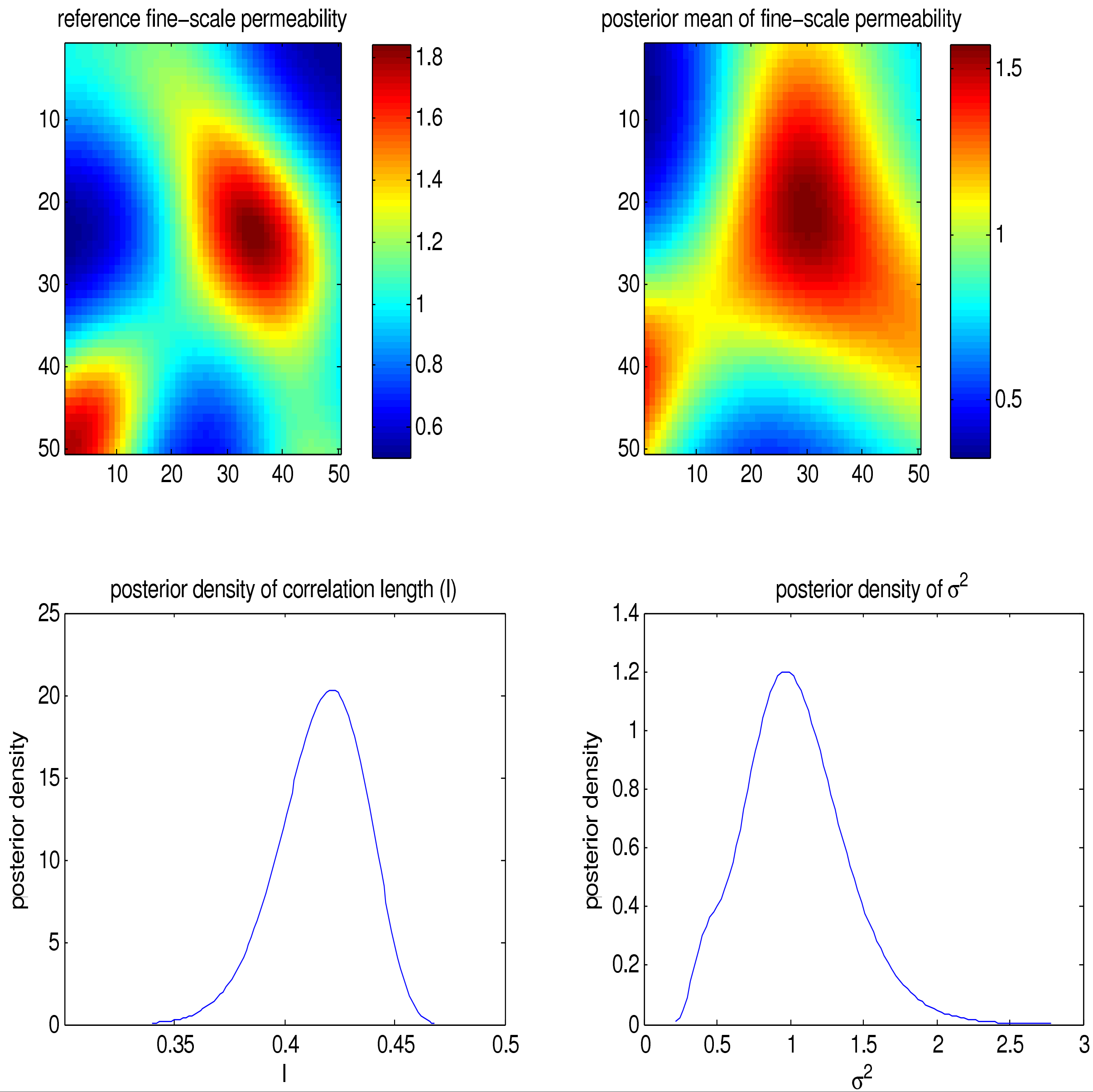

posterior density of the correlation length(l)
posterior density
l
0
5
10
15
20
25
0.1
0.15
0.2
0.25
0.3
0.35


posterior density of σ2
posterior density
σ2
0
0.5
1
1.5
2
3
4
5

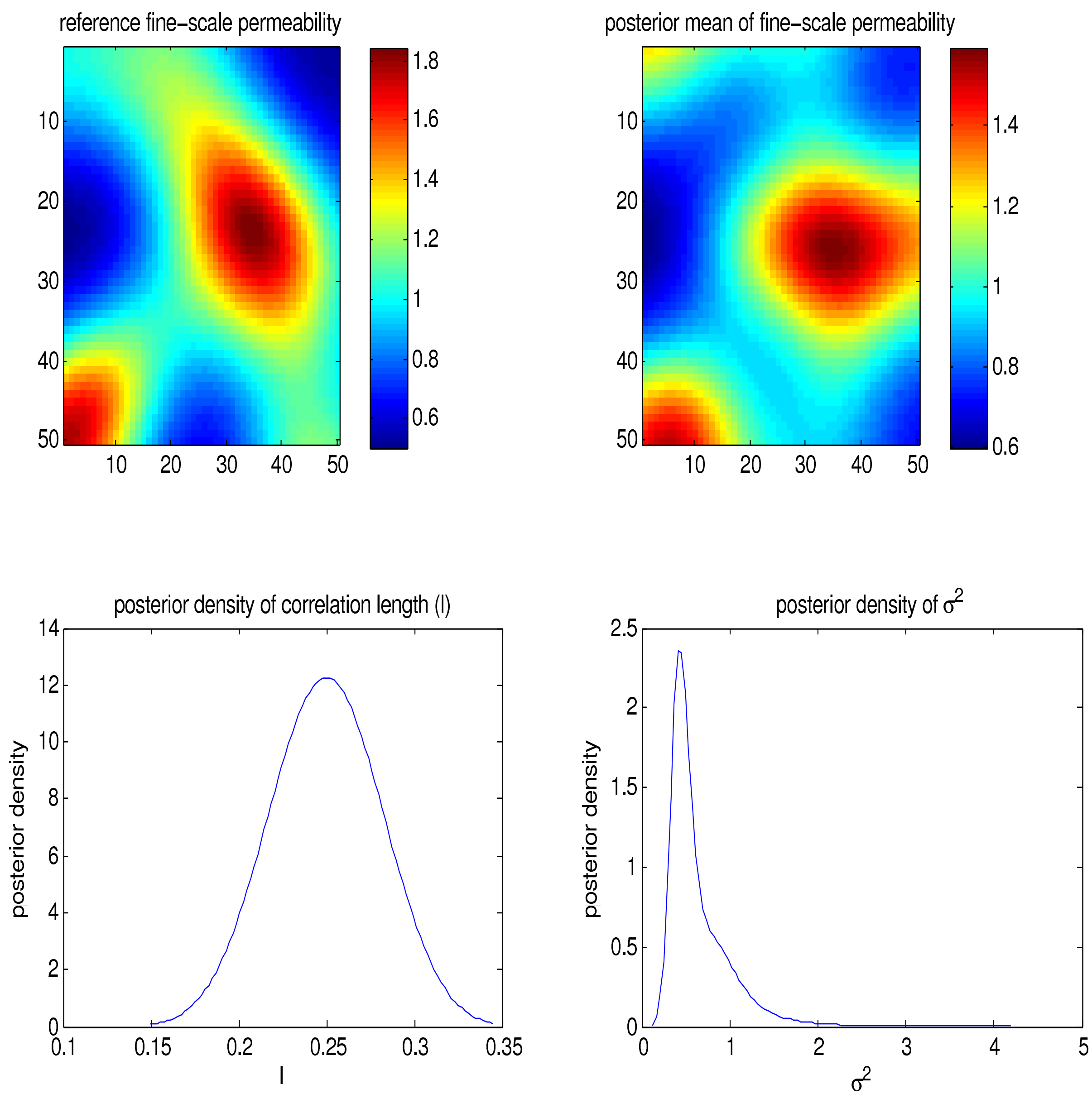

posterior density of θ0
posterior density
θ0
0
0.5
1
1.5
2
2.5
3
−1
−0.5


reference fine-scale permeability field
posterior 0.25 quantile
posterior median
posterior 0.75 quantile
10
20
30
40
50
0.6
0.8
1
1.2
1.4
1.6
1.8

10 percent fine-scale data observed and no coarse-scale data available



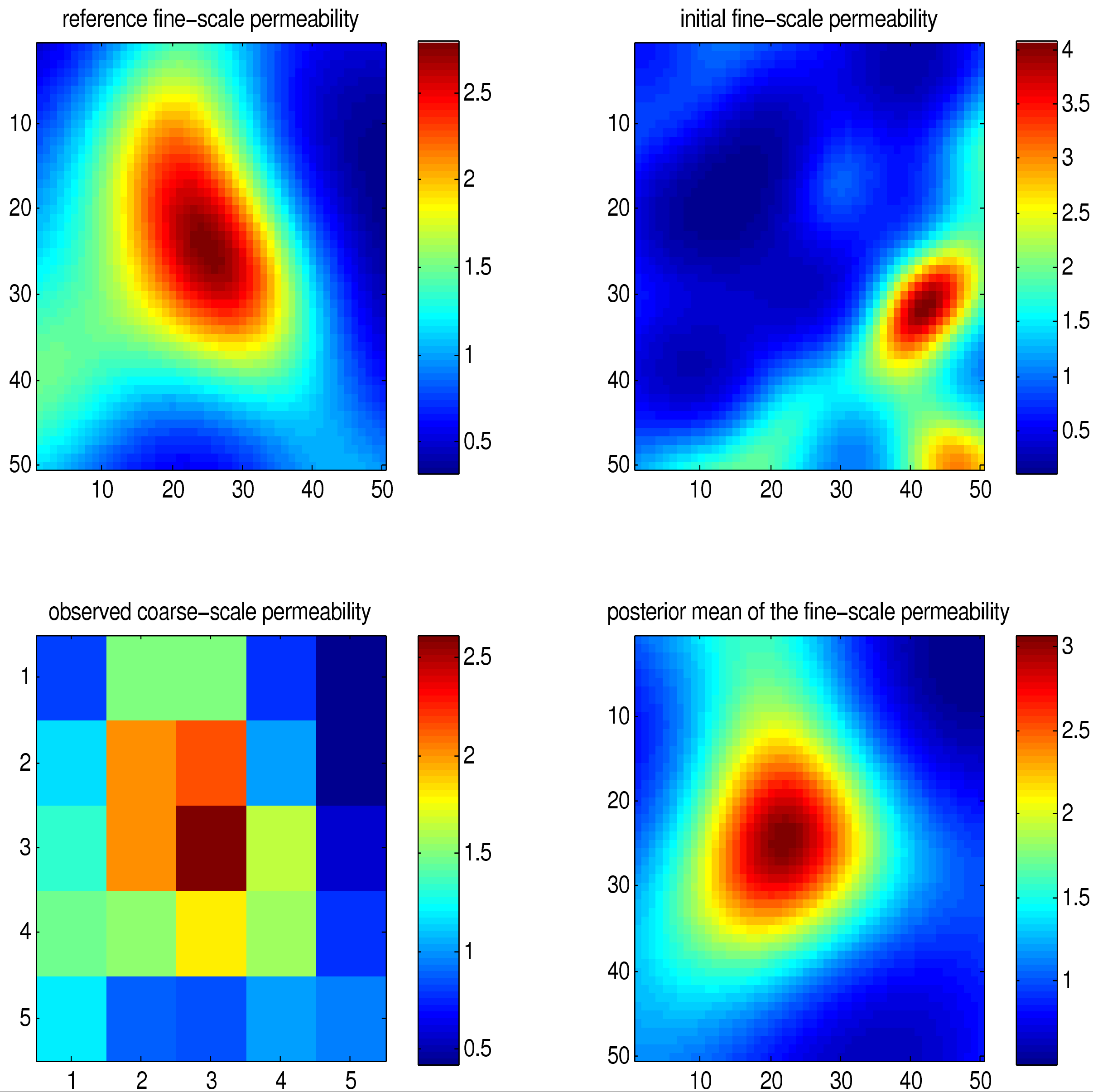
25 percent fine-scale data observed and no coarse-scale data available



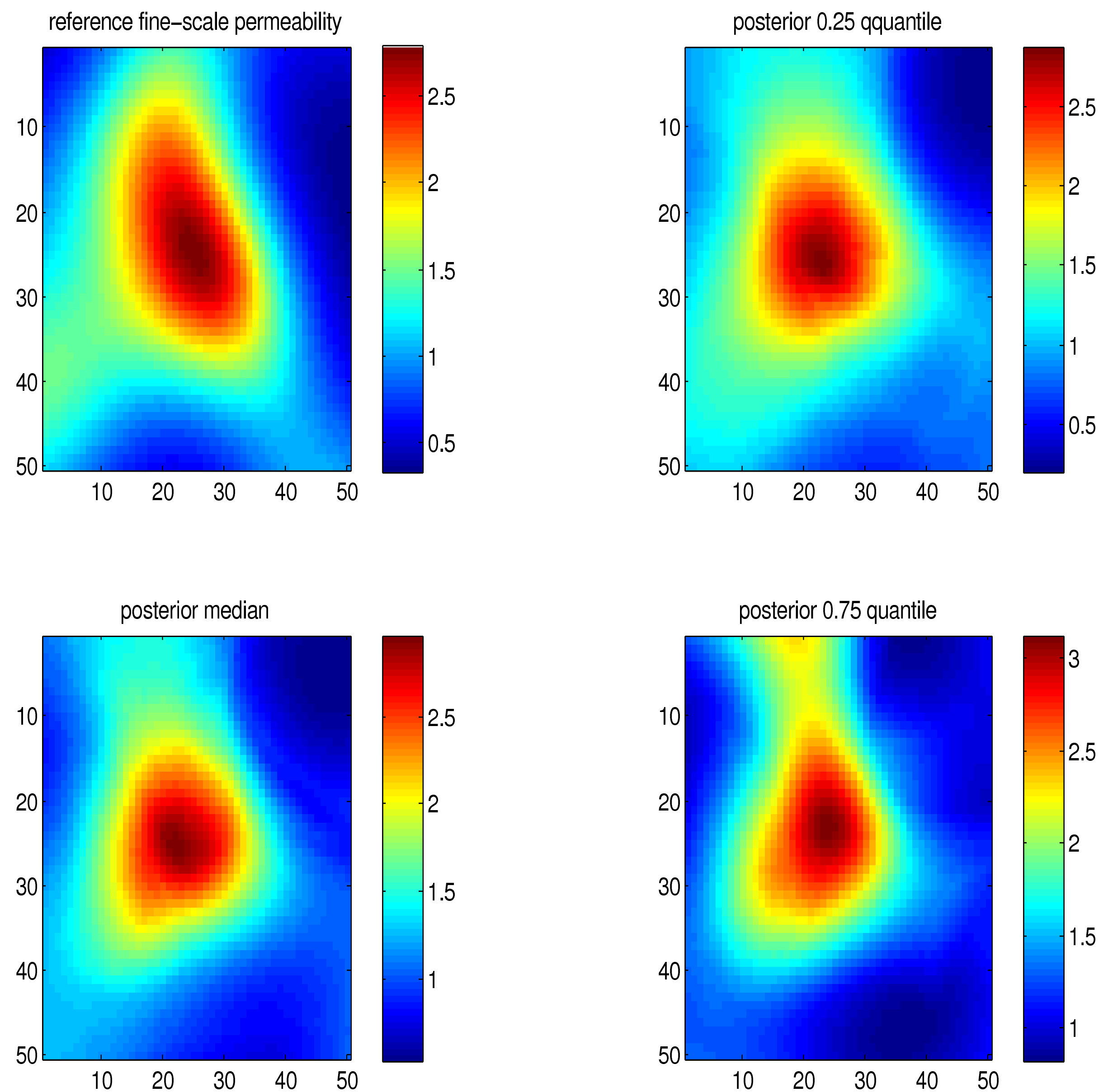
Numerical results with unknown K-L terms

- We generate 15 fine-scale permeability field with $l = .3$, $\sigma^2 = .2$ and the reference permeability field is taken to be the average of these 15 permeability field.
- We take the first 20 terms in the K-L expansion while generating the reference field.
- The mode of the posterior distribution of m comes out to be 19.
- The posterior mean of fine-scale permeability field resembles very close to the reference permeability field.
- The posterior density of l is bimodal but the highest peak is near$.3$.
- The posterior density σ^2 are centered around $.2$.

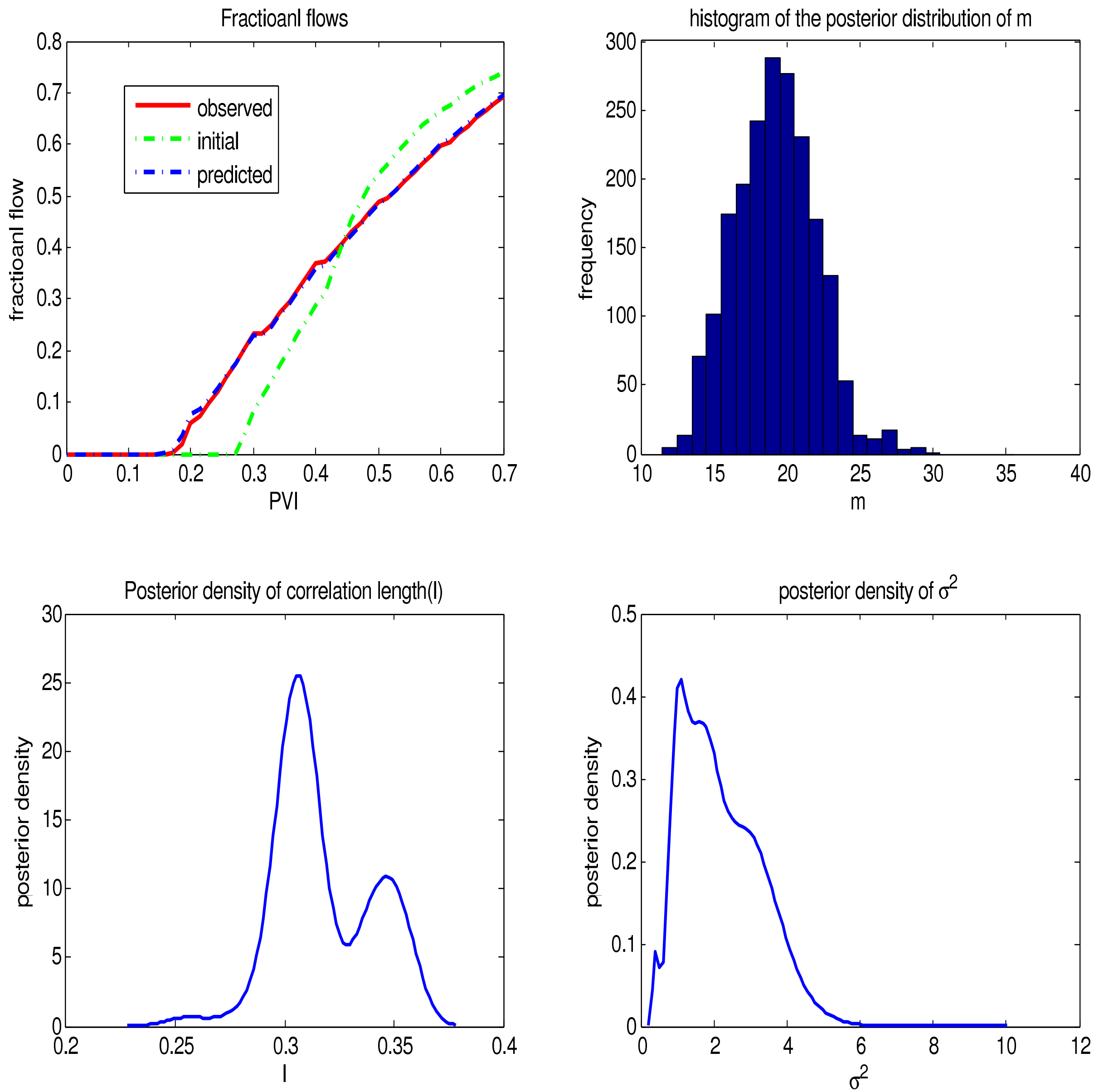
Numerical Results using Reversible Jump MCMC



Numerical Results using Reversible Jump MCMC



Numerical Results using Reversible Jump MCMC



STATISTICS
TEXAS A&M UNIVERSITY

Conclusion and future work

- Our hierarchical model is very flexible.
- If the coarse-scale data is available (even if in a very large coarse grid) our hierarchical model can efficiently quantify and reduce the uncertainty in the parameters that defines the permeability field.
- If the coarse-scale data is not available, our hierarchical model still works but then at least 25 percent of the data in fine-scale should be known.